

ICOADS Release 2.4 (1784-May 2007) IMMA Status

S. Woodruff and S. Lubker, rev. 29 August 2007

Table 1. Status of data currently used in ICOADS from the real-time (RT) and delayed-mode (DM) archives. Record sizes in bytes, and total sizes (10^9 bytes) uncompressed (Z compressed).

Archive	Period	Reports	Structure	Rec. size	Total (comp.) size
NCDC	~2003→	n/a	core+c1+c2+c3+c6	~451 ¹	n/a
ICOADS.RT	2005-May 2007	38,552,109	core+c1[+c2+c4]+c6	~324 ¹	12.5 (2.5) GB
ICOADS.DM	1784-2004	199,231,952	core+c1[+c2+c4+c6]	~273 ²	54.5 (12.1) GB
Total	1784-May 2007	237,784,061			67.1 (14.7) GB

1. Average record size. The supplementary (e.g., original GTS message) data (c6) may be variable-length (NCDC average report-length from Feb. 2004 GTS data). Differences also exist in the amounts of original GTS data and bulletin header information retained for NCDC vs. RT.

2. Average record-sizes; brackets indicate that c2, c4, and c6 were attached only if they contained extant data/metadata (thus far, c4 metadata have been attached only back to 1973).

Table 2. Sizes of IMMA format components: core and attachments (atmm) (c5 is still under development). Actual records sizes may be smaller (e.g., for 1784-2004 in Table 1, the average size of ~273B is less than nominal size of core+c1+c2+c4+c6 = 306B) because of the omission of empty attachments, and because any trailing blanks are omitted at the end of the last attachment.

Abbrev.	Name	Size (B)	Comments
core	core	108	
c1	ICOADS atmm	65	With "Adaptive" QC flags for 1784-1997 SST
c2	IMMT-2/FM 13 atmm	76	
c3	Model quality control atmm	66	VOSClm data so far unavailable in ICOADS
c4	Ship metadata atmm	57	From WMO–No. 47 for 1973-2006
c5	Historical atmm	(to be decided)	
c6	Supplemental data atmm	(variable)	

Background notes:

- a) Release numbering, currently at 2.4, designates the currently available data (and monthly summaries), drawn from the two separate ICOADS.RT and ICOADS.DM archives.
- b) Structure: The attachment structure of IMMA, and Fortran read software, was designed with the capability to save space through the omission of attachments (e.g., those not relevant for a given input dataset; e.g., Table 1). Table 2 provides sizes of the IMMA format components.
- c) NCDC marine data currently use the fixed "VOSClm" format listed in Table 1. In contrast, we have not yet included c3 for ICOADS (because c3 only derives from the UK Met Office GTS data available only for the VOSClm Project). If all of the ICOADS data were written out in the VOSClm format: Total reports from Table 1 \times 451B = 107GB (59% larger than 67.1GB).
- d) Additional NCDC format characteristics at variance with Boulder ICOADS characteristics:
 - i) Date and time with leading zeros (e.g., "200707010000"). This contrasts with the otherwise uniform numeric format model of IMMA, which has no leading zeros (e.g., "2007 7 1 0").
 - ii) Longitude expressed according to the NCDC convention (–179.99-180.00E) as opposed to the regular ICOADS convention (0.00-359.99E).
- e) ICOAD.RT: Converted from NCEP "BUFR+string." Less fully processed than ICOADS.DM. Data through Sep. 1999 were converted from the attached original GTS message strings, to resolve some early BUFR problems. Data from Oct. 1999-May 2007 were processed from BUFR into IMMA. Note: additional ICOADS.RT data, overlapping ICOADS.DM, are available back to March 1997 (not listed in Table 1).
- f) ICOADS.DM: For Release 2.4, the 1784-1997 IMMA data were re-done using the binary LMR format in order to merge important supplementary data (into c6). As resources permit, those supplementary data should be tapped for regular fields not previously defined in ICOADS formats (LMRF/LMR) but now available in IMMA (e.g., sea ice fields), or planned for availability in IMMA when the historical atmm is finalized (e.g., Beaufort wind force numbers).
- g) Dave Berry (UK NOCS) provided metadata records selected from WMO–No. 47 (ship characteristics), which were matched one-for-one with IMMA.